On Schwarzschild causality in higher dimensions

Peter Cameron[*] and **Maciej Dunajski**

Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, United Kingdom

E-mail: pjc96@cam.ac.uk and m.dunajski@damtp.cam.ac.uk

Abstract

We show that the causal properties of asymptotically flat spacetimes depend on their dimensionality: while the time-like future of any point in the past conformal infinity $\mathcal{I}^-$ contains the whole of the future conformal infinity $\mathcal{I}^+$ in $(2+1)$ and $(3+1)$ dimensional Schwarzschild spacetimes, this property (which we call the Penrose property) does not hold for $(d+1)$ dimensional Schwarzschild if $d > 3$. We also show that the Penrose property holds for the Kerr solution in $(3+1)$ dimensions, and discuss the connection with scattering theory in the presence of positive mass.

Keywords: conformal infinity, causality, Schwarzschild solution

(Some figures may appear in colour only in the online journal)

1. Introduction

In his criticism of a theory of quantum gravity, Penrose [11] has considered the problem of whether any pair of endless timelike curves in an asymptotically flat Lorentzian spacetime can be connected by a timelike curve (figure 1). This causal property is conformally invariant, and can be reformulated (theorem 2.5) in terms of past and future conformal infinities $\mathcal{I}^-$ and $\mathcal{I}^+$. We shall say that a weakly asymptotically simple Lorentzian manifold (see section 2 for precise assumptions and definitions) admits the Penrose property if and only if the timelike future of any point $p \in \mathcal{I}^-$ contains the whole of $\mathcal{I}^+$ (see also [17]). Penrose established that while the property does not hold for Minkowski space (in any dimension), it does in fact hold for positive mass Schwarzschild spacetime in $3+1$ dimensions. We will consider the Penrose property in

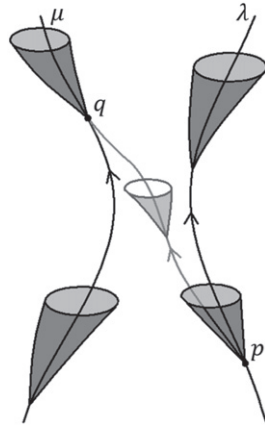[*]Author to whom any correspondence should be addressed.

Figure 1. Two endless timelike curves connected by a timelike curve.

a more general setting and investigate how it is affected by dimension and by the presence of both positive and negative mass. The results can be summarised by the following theorem:

Theorem A.   *The Penrose property is satisfied by Schwarzschild spacetime of mass m and varying spacetime dimension according to the following table*

Spacetime dimension	$m > 0$	$m \leqslant 0$
3	✓	✗
4	✓	✗
$\geqslant 5$	✗	✗

Besides the quantum gravity motivation of [11], we also seek to clear up some apparent contradictions in the literature. For example, in a footnote in section 2 of [19], Witten states that one can leave the dimension arbitrary when studying the causal properties of spacetimes and that restricting to 4 spacetime dimensions 'does not introduce any complications'. This is in contradiction with [3], where in section 2.1 the authors outline an argument which suggests that the causal properties of higher dimensional asymptotically flat spacetimes are different to those in 4 dimensions. The results of theorem A agree that the causal structure of spacetime is affected by dimensionality.

We begin in section 2 by summarising Penrose's motivation for studying the Penrose property. This comes from considering, as he describes it, a 'Lorentz covariant' approach to quantum gravity. In section 3 we consider Minkowski spacetime in more detail and determine which pairs of points on $\mathcal{I}^-$ and $\mathcal{I}^+$ can be connected by timelike curves. In section 4.1, we then consider the effect of inserting a point mass in $2 + 1$ dimensions. We will see that only in the case of positive mass is the Penrose property satisfied. In section 4.2.1 we review Penrose's argument that the Penrose property does hold for Schwarzschild in $3 + 1$ dimensions. By comparison with previous results for Minkowski spacetime, we then show in section 4.2.2 that this is not true in the negative mass case. In sections 4.3.1 and 4.3.2 we investigate the higher dimensional Schwarzschild metrics of positive and negative mass respectively. The results are summarised again in section 4.4 and we observe that the Penrose property appears

to be a property of positive mass spacetimes in low dimensions. In section 5 we show how the methods used in previous sections can be extended to study the Penrose property for the (positive mass) Kerr metric in $3 + 1$ dimensions. This case is more complicated because the metric is no longer spherically symmetric. Nevertheless, we are able to prove the following theorem:

Theorem B. *The (positive mass) Kerr spacetime in $3 + 1$ dimensions satisfies the Penrose property.*

Finally, in section 6, we show that the results found previously are consistent with the recent work of [1] on unique continuation from null infinity of the linear wave equation in various spacetimes. We also discuss the Penrose property for general static, spherically symmetric spacetimes.

2. 'Lorentz covariant' quantum gravity

In [11], Penrose considers a 'Lorentz covariant' approach to quantum gravity where the theory is constructed with respect to some background Minkowski metric η. The physical field tensors, including the metric g, are constructed as an infinite sum of tensors defined with respect to this background metric. For example, we write a physical operator as a sum

$$\mathcal{O}_g(x^\mu) = \mathcal{O}_\eta^{(0)}(x^\mu) + \mathcal{O}_\eta^{(1)}(x^\mu) + \mathcal{O}_\eta^{(2)}(x^\mu) + \ldots \tag{2.1}$$

where the operators $\mathcal{O}_\eta^{(\alpha)}$ are Lorentz covariant operators such that for any states $|\psi\rangle$, $|\psi'\rangle$ and any point x^μ, we have $\left|\langle\psi'|\mathcal{O}_\eta^{(\alpha)}(x^\mu)|\psi\rangle\right| \to 0$ sufficiently quickly as $\alpha \to \infty$ so that this sum converges.

This construction places restrictions on the physical theory one can write down. In particular, the physical metric g defines the causal structure of the theory and we need to ensure this causal structure is consistent with the standard rules of quantum field theory. For example, we know that any two operators evaluated at spacelike separated points must commute, while operators evaluated at timelike separated points should have non-zero commutator in general (where we evaluate this condition using η for the background operators and g for the physical operators). This leads to a consistency condition on the theory. In order for the 'Lorentz covariant' approach to hold, we will require that the lightcones defined by g should be contained inside the lightcones defined by η. Following [11], we will denote this condition[1] by $g < \eta$, since it tells us that if a curve is timelike with respect to g, then it must also be timelike with respect to η. If the condition $g < \eta$ fails, then there are points a, b as shown on the right in figure 2 which are timelike separated with respect to g but spacelike separated with respect to η. This means we require the background Lorentz covariant operators defined at these points to commute, but the physical operators should have non-zero commutator in general. This is clearly incompatible with the Lorentz covariant expansion (2.1).

We now make some definitions which will be used throughout this paper.

Definition 2.1. A *spacetime* (M, g) is a connected manifold M of dimension $d + 1$, where $d \geqslant 1$, equipped with a Lorentzian metric g of signature $(1, d)$. We assume also that the metric tensor, g_{ab}, is C^3 since this will be required for theorem 4.6.

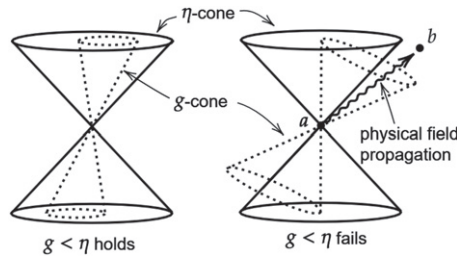[1] We are using the 'mostly minus' signature $(+, -, \ldots, -)$.

Figure 2. $g < \eta$ is a necessary condition for the 'Lorentz covariant' approach to be consistent with QFT.

Following [9], we make some further definitions

Definition 2.2. A time- and space-orientable spacetime (M, g) is *asymptotically empty and simple* if there exists a strongly causal spacetime $(\tilde{M}, \tilde{g})$ and an embedding map $\lambda : M \to \tilde{M}$ such that

(a) $\tilde{M} = \lambda(M) \cup \partial \tilde{M}$, where $\partial \tilde{M}$ is the boundary of $\tilde{M}$;
(b) $\lambda^*(\tilde{g}) = \Omega^2 g$, where Ω is a smooth function on $\tilde{M}$ such that $\Omega > 0$ on $\lambda(M)$ and $\Omega = 0$, $d\Omega \neq 0$ on $\partial \tilde{M}$
(c) Every null geodesic in (M, g) is mapped by λ to a null geodesic in $(\tilde{M}, \tilde{g})$ which has two endpoints on $\partial \tilde{M}$;
(d) The Ricci tensor of g vanishes on an open neighbourhood of $\partial \tilde{M}$ in $\tilde{M}$.

We will mostly omit the embedding map λ since the identification between M and $\tilde{M}$ should be clear (the exceptions to this are in sections 4.3.2 and 5). Definition 2.2 ensures that a conformal compactification $(\tilde{M}, \tilde{g})$ exists for this spacetime. Condition (c) however is extremely restrictive, since it rules out spacetimes containing black hole regions. We will instead consider spacetimes which are *weakly asymptotically empty and simple*.

Definition 2.3. A spacetime (M, g) is *weakly asymptotically empty and simple* if there is an asymptotically empty and simple spacetime (M', g') and a neighbourhood U' of $\partial \tilde{M}$ in M' such that $U' \cap M'$ is isometric to an open set U of M.

We will also restrict ourselves to spacetimes containing only one asymptotically flat end. The boundary of the conformally related spacetime $(\tilde{M}, \tilde{g})$ can then be split into five parts in the usual way [18]—two null surfaces $\mathcal{I}^-$ and $\mathcal{I}^+$ (known as past and future null infinity respectively), spatial infinity denoted i^0 as well as past and future timelike infinity denoted i^- and i^+ respectively. All null geodesics which do not enter a black hole region will have past endpoint on $\mathcal{I}^-$ and future endpoint on $\mathcal{I}^+$. Similarly, all timelike geodesics which do not enter a black hole region will have past endpoint on i^- and future endpoint on i^+. The definition of a weakly asymptotically empty and simple spacetime includes the Minkowski, Schwarzschild and Kerr spacetimes.

We will also use units in which $c = G = 1$.

In order to understand the condition $g < \eta$ on a spacetime, (M, g), we shall consider the following property:

Definition 2.4. A spacetime, (M, g), with conformal compactification $(\tilde{M}, \tilde{g})$ has the *Penrose property* if any points $p \in \mathcal{I}^-$ and $q \in \mathcal{I}^+$ can be timelike connected (i.e. connected by a smooth timelike curve).
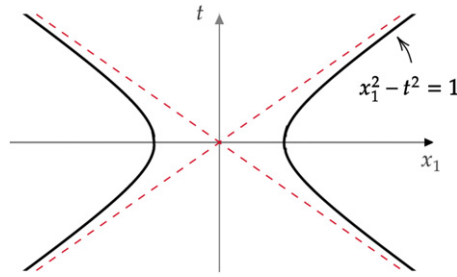
Figure 3. Two everywhere spacelike-separated branches of the hyperbola $x_1^2 - t^2 = 1$. These branches tend towards null lines as we approach infinity.

An equivalent definition given in [11] which makes no mention of conformal compactification is a consequence of the following:

Theorem 2.5 (Penrose [11], statement that theorem 4.3 is equivalent to theorem 4.4). *A spacetime, (M, g), has the Penrose property if and only if given any two endless timelike curves λ and μ in M, there exist points $p \in \lambda$, $q \in \mu$ and a future directed timelike curve from p to q.*

To begin studying the Penrose property it is helpful to consider the simple example of Minkowski spacetime in $d + 1$ dimensions ($d \geqslant 2$):

Theorem 2.6 (Penrose [11], theorem 4.5). *The Penrose property is not satisfied by Minkowski spacetime in $d + 1$ dimensions, $\mathbb{M}^{(1,d)}$, for any $d \geqslant 1$.*

Proof. With respect to co-ordinates $(t, x_1, x_2, \ldots, x_d)$, the Minkowski line element in $d + 1$ dimensions can be written as

$$\mathrm{d}s^2 = \mathrm{d}t^2 - \mathrm{d}x_1^2 - \mathrm{d}x_2^2 - \cdots - \mathrm{d}x_d^2 \tag{2.2}$$

Now consider the hyperbola in the $x_1 - t$ plane defined by $x_1^2 - t^2 = 1$, $x_2 = \cdots = x_d = 0$. We see (figure 3) that the two branches of this hyperbola are endless timelike curves which are everywhere spacelike separated. We therefore conclude, by theorem 2.5, that the Penrose property does not hold for Minkowski spacetime in any dimension. $\square$

Penrose's motivation for studying the Penrose property in [11] is that it can be related to the consistency condition $g < \eta$ established previously for the 'Lorentz covariant' quantum gravity construction:

Proposition 2.7 (Penrose [11], statement that theorem 4.5 follows from theorems 4.3 or 4.4). *If the spacetime (M, g) satisfies the Penrose property, then the condition $g < \eta$ fails.*

In the proof of this proposition, we assume that the compactified manifolds of the physical and background spacetimes are diffeomorphic, at least in some neighbourhood of $\mathcal{I}^- \cup i^0 \cup \mathcal{I}^+$ (see definition 2.3). We expect this to be the case since g and η should have the same asymptotics if the 'Lorentz covariant' theory is to be able to describe physical situations such as scattering experiments.

Proof. If $g < \eta$ is satisfied, then curves which are timelike with respect to the physical metric, g, must also be timelike with respect to the background metric, η (and this is also true for
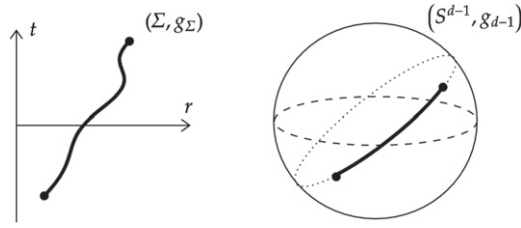
Figure 4. A curve in $d + 1$-dimensional, spherically symmetric, static spacetime can be thought of as a pair of curves in Σ and S^{d-1}.

their respective compactifications). So if the Penrose property were to hold for g then it would also hold for η—we can use exactly the same path through the compactified manifolds. But we have seen that the Penrose property does not hold for Minkowski space which leads to a contradiction and we conclude that g does not satisfy the Penrose property. $\square$

We will consider the Penrose property as a property in its own right, rather than as a tool to constrain possible theories of quantum gravity. It will turn out that a useful way to prove results relating to this property will be to use a similar trick of comparing the metric of interest to some other metric, defined on the same manifold, whose properties we do understand.

3. Minkowski causality

Before looking at more complicated situations, we should first make sure we fully understand Minkowski spacetime. We have seen from the hyperbola example that in compactified Minkowski space, some points on $\mathcal{I}^-$ cannot be timelike connected to certain points on $\mathcal{I}^+$. To understand exactly which points this applies to, we first note the following lemmas:

Lemma 3.1. *Consider a static, spherically symmetric spacetime, (M, g), which is weakly asymptotically flat at null infinity. Let γ be a timelike curve with endpoints $p \in \mathcal{I}^-$ and $q \in \mathcal{I}^+$. Then we can choose polar co-ordinates $(t, r, \theta_1, \ldots, \theta_{d-2}, \phi)$, where $t \in \mathbb{R}$, r is defined on some subset of $\mathbb{R}$, $\phi \in [0, 2\pi)$ and $\theta_1, \ldots \theta_{d-2} \in [0, \pi]$ as well as a timelike curve, γ', from p to q such that $\theta_1 = \ldots = \theta_{d-2} = \pi/2$ along γ'.*

In other words, it suffices to look for timelike curves between p and q which lie in a submanifold $\Sigma \times S^1 \subset M$, where Σ a two-dimensional surface with co-ordinates (t, r).

Proof.   Let the metric take the form

$$g = A(r)^2 \mathrm{d}t^2 - B(r)^2 \mathrm{d}r^2 - r^2 g_{d-1},$$

where g_{d-1} is the round metric on a unit $(d - 1)$–dimensional sphere S^{d-1}. We follow [10] and consider the conformally re-scaled metric

$$\bar{g} = r^{-2}g = g_\Sigma - g_{d-1},$$

defined on the product space $\Sigma \times S^{d-1}$, where g_Σ is a Lorentzian metric on the surface Σ. An endless non–spacelike curve in the $d + 1$ dimensional spacetime therefore corresponds to a pair of curves described by a common parameter, s, which we choose to be the arc length of the second curve (figure 4). One of these is an endless curve in Σ and the other is a curve in
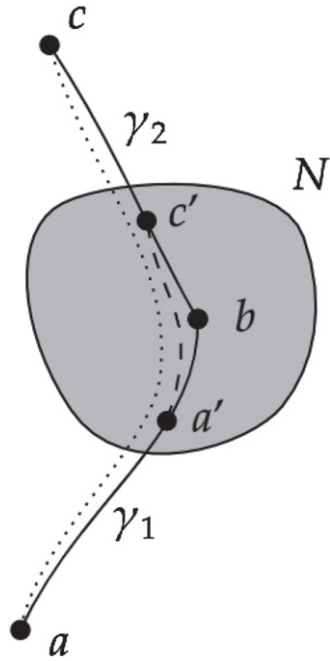
Figure 5. Assume $\gamma_1 \cup \gamma_2$ is not a single null geodesic. Then we choose the points a' and c' to lie in a normal neighbourhood of b, so there exists a timelike geodesic between them (dashed line). Thus we have connected a to c using three causal geodesic segments, where in particular one of these segments is timelike. This can be smoothed to give a timelike curve from a to c (dotted line).

S^{d-1} (which may be a single point) with endpoints which come from looking at the points on S^{d-1} at which γ meets $\mathcal{I}^{\pm}$ [2].

If the full curve in M is timelike, then the line element evaluated at each point along it satisfies

$$\bar{ds}^2 = ds_\Sigma^2 - d\omega_{d-1}^2 > 0$$

where ds_Σ^2 and $d\omega_{d-1}^2$ are the line elements along the curves in (Σ, g_Σ) and (S^{d-1}, g_{d-1}) respectively. Now consider the curve in (S^{d-1}, g_{d-1}). The metric g_{d-1} is a Riemannian metric, so $d\omega_{d-1}^2(s) \geqslant 0$ for each value of the parameter s (with equality corresponding to a radial line segment in M). We can modify this curve, keeping the endpoints the same, so that it becomes a geodesic. Since we chose s to be the arc length, this reduces the value of $d\omega_{d-1}^2(s)$ at each value of s and ensures that the curve in the full spacetime remains timelike. Furthermore, the geodesic in S^{d-1} is a segment of a great circle, so we can choose spherical co-ordinates $(\theta_1, \ldots, \theta_{d-2}, \phi)$ (where $\theta_1, \ldots, \theta_{d-2} \in [0, \pi]$, $\phi \in [0, 2\pi)$) such that $\theta_1 = \cdots = \theta_{d-2} = \pi/2$ along the geodesic. This means that the curve in the full spacetime is restricted to the surface $\theta_1 = \ldots = \theta_{d-2} = \pi/2$, i.e. a $\Sigma \times S^1$ submanifold of M. $\qquad\square$

[2] Recall that both $\mathcal{I}^+$ and $\mathcal{I}^-$ are topologically $\mathbb{R} \times S^{d-1}$—this is a straightforward extension of proposition 6.9.4 in [9]).

Proposition 3.2 (Penrose [12], proposition 2.20 and proposition 2.23). *Let γ_1 and γ_2 be future pointing causal (i.e. timelike or null) geodesics from a to b and from b to c respectively. Then either there exists a smooth timelike curve from a to c or the union of these geodesics, $\gamma = \gamma_1 \cup \gamma_2$, is itself a null geodesic (figure 5).*

We now return to considering Minkowski spacetime in $d + 1$ dimensions, $\mathbb{M}^{(1,d)}$. Since this spacetime is spherically symmetric, proposition 3.2 ensures that if two points can be connected by a timelike curve then we can choose this curve to lie in a three-dimensional surface with induced metric equal to that of $\mathbb{M}^{(1,2)}$. It therefore suffices to consider only the compactification of Minkowski spacetime in $2 + 1$ dimensions. The Minkowski line element in $2 + 1$ dimensions in polar co-ordinates is

$$\mathrm{d}s^2 = \mathrm{d}t^2 - \mathrm{d}r^2 - r^2 \ \mathrm{d}\phi^2 \tag{3.3}$$

We define retarded and advanced time co-ordinates respectively by

$$u = t - r, \quad v = t + r \tag{3.4}$$

We then compactify the metric by defining

$$u = \tan \ P, \quad v = \tan \ Q \tag{3.5}$$

Finally, we define new co-ordinates

$$T = Q + P \in (-\pi, \pi), \quad \chi = Q - P \in [0, \pi) \tag{3.6}$$

which we can think of as 'time' and 'radial' co-ordinates respectively in the compactified spacetime.

We then consider the conformally related metric $\tilde{g}$, given by

$$\tilde{g} = \Omega^2 g = (2 \ \cos \ P \ \cos \ Q)^2 g$$

This gives the line element

$$\tilde{\mathrm{d}s}^2 = \mathrm{d}T^2 - \mathrm{d}\chi^2 - \sin^2 \chi \ \mathrm{d}\phi^2 \tag{3.7}$$

In $2 + 1$ dimensions, the compactified manifold, $\tilde{M}$, is a double cone, as shown in figure 6, where we can think of time as moving upwards. The antipodal points p and q in figure 6 correspond to the past and future endpoints of different branches of the hyperbola shown in figure 3. They have T and χ co-ordinates $(T, \chi) = (-\frac{\pi}{2}, \frac{\pi}{2}), (\frac{\pi}{2}, \frac{\pi}{2})$ respectively. The radial null geodesic between these points is a straight line in this diagram. Moreover, any null geodesic from p which enters the interior spacetime, M, has future endpoint at q. For a general point $p' \in \mathcal{I}^-$, there is a unique $q' \in \mathcal{I}^+$ which is the future endpoint of null geodesics from p' which enter M. The points p' and q' can be obtained from p and q by a simple translation in T and ϕ.

Proposition 3.3. *Let $p \in \mathcal{I}^-$ and let q be the unique point in $\mathcal{I}^+$ which is connected to p by a null geodesic which enters M. Then the only points in $\mathcal{I}^+$ which cannot be timelike connected to p are the points in $\mathcal{I}^+$ which can be reached from q by a past pointing null geodesic, including the point q itself. That is, for any $p \in \mathcal{I}^-$*

$$I^+(p) \cap \mathcal{I}^+ = \mathcal{I}^+ \backslash \left(J^-(q) \cap \mathcal{I}^+ \right) \tag{3.8}$$
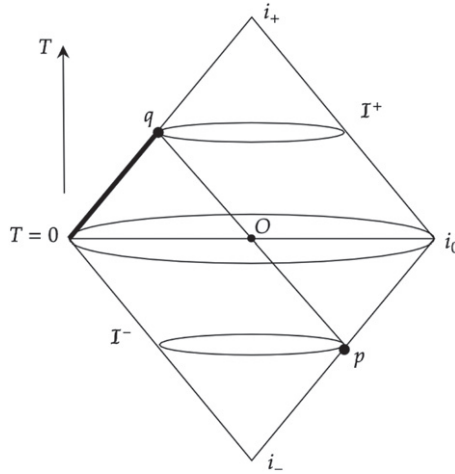
Figure 6. The points $p \in \mathcal{I}^-$ and $q \in \mathcal{I}^+$ are connected by a radial null geodesic through the origin. The point p can be timelike connected to all points in $\mathcal{I}^+$ except those that lie to the past of q and on the same null generator of $\mathcal{I}^+$. These points correspond to the shaded interval.

where $q \in \mathcal{I}^+ \cap \partial J^+(p)$ can be connected to p by a null geodesic which does not pass through i^0 (and hence enters M) and where we define the following sets

$$J^+(p) = \{q \in \tilde{M} | \exists \, a \, smooth \, future\text{-}directed \, causal \, curve \, from \, p \, to \, q\}$$

$$J^-(p) = \{q \in \tilde{M} | p \in J^+(q)\}$$

$$I^+(p) = \{q \in \tilde{M} | \exists \, a \, smooth \, future\text{-}directed \, timelike \, curve \, from \, p \, to \, q\}$$ (3.9)

$$I^-(p) = \{q \in \tilde{M} | p \in I^+(q)\}$$

For non-compact spacetimes, (M, g), we can define the same sets as subsets of M rather than as subsets of $\tilde{M}$.

This proposition is illustrated in figure 7. In particular, we note that antipodal points close to spacelike infinity cannot be timelike connected.

Proof.   Let $p \in \mathcal{I}^-$ have co-ordinates $(T, \chi, \phi) = (-\frac{\pi}{2}, \frac{\pi}{2}, 0)$, so q has co-ordinates $(\frac{\pi}{2}, \frac{\pi}{2}, \pi)$. This is shown in figure 6. These points correspond to the past and future endpoints of separate branches of the hyperbola shown in figure 3. A general point $p' \in \mathcal{I}^-$ is related to p by a translation in T and ϕ. Since the compactified metric (3.7) does not depend on these co-ordinates, it suffices to prove the proposition for this point p. The same argument would then apply to general $p' \in \mathcal{I}^-$, only with the point $q \in \mathcal{I}^+$ shifted in the same way to become q'. This is shown in figure 7 for a translation in T (we assume without loss of generality that $\phi = 0$ at p').

First we note that it is clear none of the shaded points in figure 6 lie in the interior of the future lightcone from p. Next observe that it is straightforward to find a timelike curve from p to points such as $\hat{q}$ in figure 8 which lie on $\mathcal{I}^+$ with T-value greater than or equal to that of q (and excluding q itself). We can use two radial null geodesics intersecting at the origin O as well as a radial null geodesic up $\mathcal{I}^+$ if necessary. Since $\hat{q} \neq q$, the total curve is not a null geodesic, so we can use proposition 3.2 to deduce that there exists a smooth timelike curve from p to $\hat{q}$. By using a curve of constant ϕ which is also a straight line in (T, χ) co-ordinates,
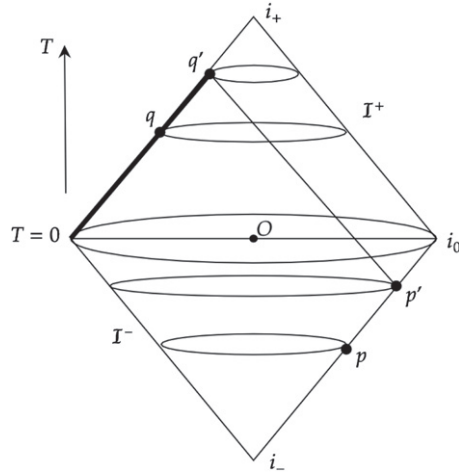
Figure 7. The point $p' \in \mathcal{I}^-$ can be connected to all points on $\mathcal{I}^+$ except those that lie to the past of q' and on the same null generator of $\mathcal{I}^+$ (i.e. the shaded interval), where q' is the unique future endpoint of null geodesics from p' which enter M.
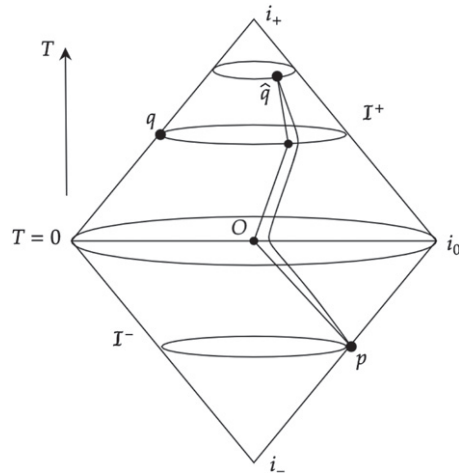

Figure 8. p can be timelike connected to any point on $\mathcal{I}^+$ with T co-ordinate greater than or equal to that of q (other than q itself). The path shown consists of two radial null lines through the origin as well as a null line up $\mathcal{I}^+$. Proposition 3.2 then implies the existence of a smooth timelike curve between p and $\hat{q}$.

we are able to timelike connect p to any point in $\mathcal{I}^+$ which has the same ϕ co-ordinate as p. This means we can concentrate on points on $\mathcal{I}^+$ with larger χ co-ordinates than p, i.e. which are closer to spacelike infinity, and whose ϕ co-ordinate differs from that of p by a value in $(0, \pi)$ (without loss of generality).

It is helpful to look at the double cone from above, as in figure 9. The circles shown correspond to surfaces $\chi = $ constant, with the larger circle $\chi = \pi$ representing spatial infinity. To show that all non-antipodal points can be timelike connected, we will follow an argument similar to the one used by Penrose in [11]. This argument, which will be used again in
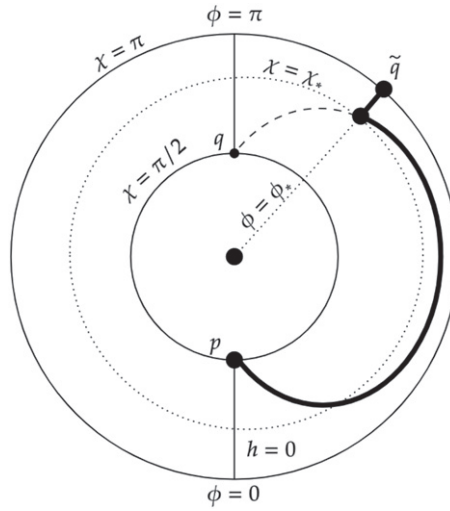
Figure 9. Double cone of compactified Minkowski spacetime viewed from above.

section 4.2.1, relies on repeated applications of proposition 3.2 to a curve which is composed of segments of null geodesics. More specifically, the curve is based on several radial null geodesics, as well as a single non-radial one, which we use to achieve the desired change in the angular co-ordinate ϕ.

From the null geodesic equations, we have

$$\dot{T} = 1 \,(\text{w log})$$

$$\sin^2(\chi)\dot{\phi} = h = \text{constant} \tag{3.10}$$

$$\dot{\chi}^2 = 1 - \frac{h^2}{\sin^2 \chi}$$

where overdots denote differentiation with respect to some parameter which we have normalised so that $\dot{T} = 1$.

We can integrate these equations to find explicit expressions for χ and ϕ involving T. Firstly we have

$$\tan^2 T = \frac{\sin^2 \chi - h^2}{\cos^2 \chi} \tag{3.11}$$

where we have chosen the geodesic to intersect $\mathcal{I}^-$ at $(T, \chi) = (-\pi/2, \pi/2)$. This ensures that the maximum value of χ is achieved at $T = 0$ and that the curve intersects $(T, \chi) = (\pi/2, \pi/2)$. Using this, we can solve for ϕ along the curve (assuming $h \neq 0$). We have

$$\sin^2 \chi \dot{\phi} = h$$

$$\implies \dot{\phi} = \frac{h\left(1 + \tan^2 T\right)}{h^2 + \tan^2 T} \tag{3.12}$$

$$\implies \phi = \tan^{-1}\left(\frac{\tan T}{h}\right) + \frac{\pi}{2}$$

Where we have chosen our geodesic to have $\phi = 0$ at $\mathcal{I}^-$. This ensures that $\phi = \pi$ on $\mathcal{I}^+$ for $h \neq 0$.

Consider a point $\tilde{q} \in \mathcal{I}^+$ at which $\chi = \chi_* \in (\pi/2, \pi)$ and $\phi = \phi_* \in (0, \pi)$. It suffices to show that a null geodesic with past endpoint at p reaches a point with these same values of the χ and ϕ co-ordinates, since we could then reach $\tilde{q}$ by following this null geodesic until we reach such a point before switching to a timelike geodesic of constant χ, ϕ. This construction is shown as the bold path in figure 9. Finally we could then apply proposition 3.2 to obtain a smooth timelike curve with the same endpoints at p and $\tilde{q}$.

Re-arranging equation (3.12), the null geodesic intersects the required point provided we choose

$$h^2 = \frac{\sin^2\chi_*}{1 + \cos^2\chi_* \tan^2(\phi_* - \pi/2)}$$

where we choose the positive sign for h since we assumed $\phi_* \in (0, \pi)$.

Choosing this value of h gives us the required null geodesic and we conclude that points in $\mathcal{I}^-$ with $\chi = T = \frac{\pi}{2}$ are timelike connected to all non-antipodal points in $\mathcal{I}^+$. □

4. The Penrose property in Schwarzschild spacetime

We now consider the effect of adding mass. In particular, we will consider the Schwarzschild spacetime in various dimensions and for both positive and negative mass.

4.1. Schwarzschild spacetime in 2 + 1 dimensions

4.1.1. Positive mass. We begin by focussing on the $2 + 1$ dimensional case, where we prove the following result:

Proposition 4.1. *The Penrose property is satisfied by the Schwarzschild spacetime of mass m in 2 + 1 dimensions if and only if m > 0.*

Proof. We first note that the failure of the Penrose property when $m = 0$ has already been established in section 3 (since the spacetime in this case is Minkowski with a point removed). We now consider the case where the mass is non-zero.

In $2 + 1$ dimensions, the symmetries of the Riemann tensor are highly constraining, and we have [16]

$$R_{abcd} = \left(R^{mn} - \frac{1}{2}Rg^{mn} \right) \epsilon_{mab}\epsilon_{ncd} \tag{4.13}$$

where ϵ_{abc} is the fully anti-symmetric tensor of rank 3 with $\epsilon_{123} = 1$.

Away from sources, the Einstein equations imply $R_{ab} = 0$ and hence spacetime is flat. This means that gravity in $2 + 1$ dimensions is different to the theory in higher dimensions. In particular, the curvature is fully determined by the Einstein equations and there is no propagating field of gravity.

The stress–energy tensor for a point source of mass m at the origin is

$$T^{00} = m\delta(\boldsymbol{r}), \quad T^{0i} = T^{ij} = 0$$

The spacetime is static and spherically symmetric, so we can write the metric in isotropic co-ordinates as

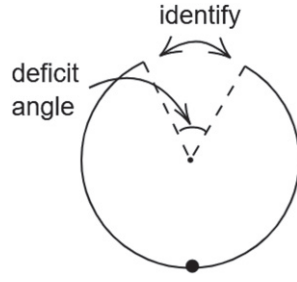$$ds^2 = N(\mathbf{r})dt^2 - \varphi(\mathbf{r})(dr^2 + r^2 \, d\phi^2)$$

Figure 10. The deficit angle resulting from inserting a positive point mass removes antipodal points.

Solving the Einstein equations in vacuum ($R_{ab} = 0$), we find ([5], equation (2.4)) that N is a constant, which we can choose to be 1 by a re-scaling of t. Next we solve for $\varphi(\mathbf{r})$ to find ([5], equation (2.6))

$$\varphi(\mathbf{r}) = r^{-8m}$$

Making the change of variables[3] $\tilde{r} = \alpha^{-1}r^{\alpha}, \tilde{\phi} = \alpha\phi$, where $\alpha = 1 - 4m$, we arrive at the metric

$$\mathrm{d}s^2 = \mathrm{d}t^2 - \mathrm{d}\tilde{r}^2 - \tilde{r}^2\,\mathrm{d}\tilde{\phi}^2 \tag{4.14}$$

We recognise this as the Minkowski metric (3.3). The important thing to note is that the co-ordinates (t, r, ϕ) do not cover all of $\mathbb{R}^3$. This is because the range of $\tilde{\phi}$ is now $0 \leqslant \tilde{\phi} < 2\pi\alpha$, so there is a conical singularity at the origin.

If $m > 0$, the range of $\tilde{\phi}$ is strictly less than 2π. We have therefore introduced a deficit angle

$$\tilde{\phi}_{\text{deficit}} = 8\pi m$$

This means that any two points on $\mathcal{I}^-$ and $\mathcal{I}^+$ will have $\tilde{\phi}$ co-ordinates which differ by strictly less than π (mod 2π), so we have effectively removed all the antipodal points (see figure 10). By the arguments presented in section 3 (in particular proposition 3.3 and its extension to general points on $\mathcal{I}^-$), this means we are able to connect any pair of points on $\mathcal{I}^-$ and $\mathcal{I}^+$ with a timelike curve, so the Penrose property is satisfied.

On the other hand, inserting a negative mass yields the opposite effect. We now have a surplus angle, so $\tilde{\phi}$ has a range strictly greater than 2π. The result of this is that given any initial point $p \in \mathcal{I}^-$, there is a wedge of points on $\mathcal{I}^+$ which cannot be reached by a timelike curve (see figure 11). At these points, the $\tilde{\phi}$ co-ordinate differs by strictly greater than π from the value at p. The Penrose property is not satisfied by this negative mass spacetime. In fact, the situation is worse than in Minkowski, since the set of points on $\mathcal{I}^+$ which cannot be reached by a timelike curve from p now forms a two-dimensional area (as opposed to the 1 dimensional line segment we found in Minkowski spacetime). $\qquad\square$

So it appears that the Penrose property is linked to the presence of positive mass[4]. This is not wholly unsurprising since we expect a positive mass source to focus and retard nearby geodesics (and to correspondingly time-advance geodesics passing far from the source as

---

[3] We are actually restricting here to $m < \frac{1}{4}$. The case $m > \frac{1}{4}$ corresponds to a 'mass at infinity', while $m = \frac{1}{4}$ gives a cylinder in the spatial dimensions. See [5] for a discussion of this. We will not consider these cases further.

[4] This link, and some of its consequences has been explored in [4, 6–8, 13].
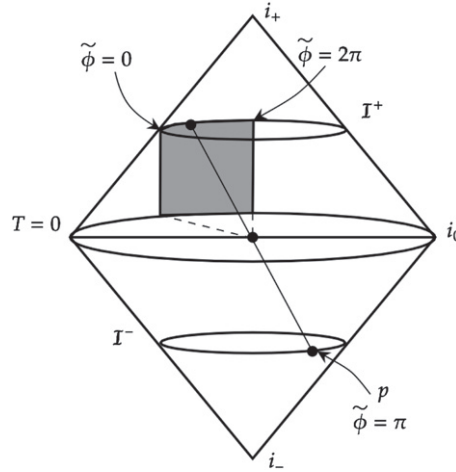
Figure 11. The surplus angle resulting from inserting a negative point mass means that given any point on $\mathcal{I}^-$ there is now a wedge of points on $\mathcal{I}^+$ which cannot be reached by a timelike curve.

described in [13]). This should make it easier for timelike curves to reach antipodal points near i^0. The results in $2+1$ dimensions, along with this intuition, suggest that we should look at positive mass spacetimes if we want to satisfy the Penrose property.

4.2. Schwarzschild spacetime in 3 + 1 dimensions

We now consider the Schwarzschild solution in $3+1$ dimensions. As outlined in theorem A, we will find that the Penrose property is satisfied when the mass is $m > 0$ but not when $m < 0$ (the case $m = 0$ corresponds to Minkowski spacetime and has already been considered in section 3).

4.2.1. Positive mass. The positive mass Schwarzschild metric in $3+1$ dimensions is

$$ds^2 = V(r)dt^2 - \frac{dr^2}{V(r)} - r^2 \left(d\theta^2 + \sin^2 \theta \ d\phi^2 \right) \tag{4.15}$$

where $V(r) = 1 - \frac{r_s}{r}$ and $r_s = 2m > 0$ (m is the mass of the spacetime).

We can conformally compactify this metric using the same procedure as was used in section 3 for Minkowski spacetime, the only difference being we now use the tortoise co-ordinate, r_*, which satisfies

$$dr_* = \frac{dr}{V(r)} \tag{4.16}$$

in order to define the retarded and advanced time co-ordinates

$$
\begin{aligned}
u &= t - r_* = t - r - r_s \ \log(r/r_s - 1) \\
v &= t + r_* = t + r + r_s \ \log(r/r_s - 1)
\end{aligned}
\tag{4.17}
$$

where we have set the arbitrary integration constant in the definition of r_* (equation (4.16)) to zero.

Just as in section 3, we define compactified 'time' and 'radial' co-ordinates $T \in (-\pi, \pi)$ and $\chi \in [0, \pi)$ to be

$$T = \arctan v + \arctan u, \quad \chi = \arctan v - \arctan u \tag{4.18}$$

The conformally compactified metric is then

$$\tilde{d}s^2 = \Omega^2 \, ds^2 = dT^2 - d\chi^2 - \frac{r^2}{V(r)r_*^2} \, \sin^2 \chi \left(d\theta^2 + \sin^2 \theta \, d\phi^2 \right), \tag{4.19}$$

where $\Omega = \frac{2 \, \cos\left(\frac{T+\chi}{2}\right) \cos\left(\frac{T-\chi}{2}\right)}{\sqrt{V(r)}}$.

The compactified manifold for the region of Schwarzschild spacetime $\{r \geqslant r_\#\}$, where $r_\#$ is the unique solution to $r_\# + r_s \log(r_\#/r_s - 1) = 0$, is the same as the compactified Minkowski manifold, namely the (higher dimensional) cone

$$\tilde{M} = \{(T, \chi) \in [-\pi, \pi] \times [0, \pi] : |T| \leqslant \pi - \chi\} \times S^2$$

Note that $r_\# > r_s$, so $V(r) > 0$ at all points in the compactified manifold.

Theorem 4.2 (Penrose [11]). *Positive mass Schwarzschild spacetime in $3 + 1$ dimensions satisfies the Penrose property.*

Proof. We follow the method used by Penrose in [11] to connect points $p \in \mathcal{I}^-$ and $q \in \mathcal{I}^+$. We also fill in some details which are omitted from [11], in particular we show how the positivity of mass ensures that we can follow null geodesics in order to reach antipodal points (see equation (4.21)).

Just as we have done in $2 + 1$ dimensions, this method uses a piecewise null geodesic construction, with the existence of a smooth timelike curve from p to q guaranteed by proposition 3.2. We choose co-ordinates (t, r, θ, ϕ) such that $\theta = \pi/2$ at p and q. The Schwarzschild metric is static and spherically symmetric, so by lemma 3.1 it suffices to consider curves restricted to the hypersurface $\theta = \pi/2$. All curves referred to in this section will be assumed to lie in this plane. Up to shifts in t and ϕ, Schwarzschild geodesics in this plane which are null and inextendible (so in particular $r \geqslant r_s$) are uniquely determined by their impact parameter, R, defined to be the minimal value of r along the curve (which we assume is obtained at $t = 0$). The geodesic Lagrangian is

$$L = V(r)\dot{t}^2 - V(r)^{-1}\dot{r}^2 - r^2\dot{\phi}^2$$

so that, in the null case,

$$L = 0, \quad h = -\frac{1}{2}\frac{\partial L}{\partial \dot{\phi}} = r^2\dot{\phi} = \text{const}, \quad \frac{1}{2}\frac{\partial L}{\partial \dot{t}} = V(r)\dot{t} = \text{const} = 1,$$

where the last equation is our choice of the parametrisation.

For such a geodesic, which in particular must have non-zero angular momentum, h, we have:

$$\begin{aligned}
\dot{r}^2 &= 1 - \left(1 - \frac{r_s}{r}\right)\frac{h^2}{r^2} \\
&\leqslant 1 - \left(1 - \frac{r_s}{R}\right)\frac{h^2}{r^2} \\
&= 1 - \frac{R^2}{r^2}
\end{aligned} \tag{4.20}$$

where we use the fact that $\dot{r} = 0$ at $r = R$ to solve for $h^2 = R^2\left(1 - \frac{r_s}{R}\right)^{-1}$.
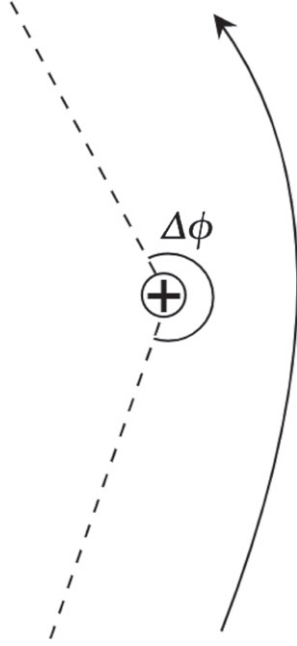
15

Figure 12. $|\Delta\phi| > \pi$ along null geodesics in positive mass Schwarzschild.

Using this, we can obtain a lower bound for $|\Delta\phi|$, the magnitude of the change in ϕ along this null geodesic:

$$\dot{\phi} = \frac{h}{r^2}$$

$$\implies |\Delta\phi| = 2R\left(1 - \frac{r_{\rm s}}{R}\right)^{-1/2} \int_R^\infty \frac{{\rm d}r}{\dot{r}r^2}$$

$$\geqslant 2R\left(1 - \frac{r_{\rm s}}{R}\right)^{-1/2} \int_R^\infty \frac{{\rm d}r}{\left(1 - R^2/r^2\right)^{1/2}r^2}$$

$$= \left[-2\left(1 - \frac{r_{\rm s}}{R}\right)^{-1/2} \tan^{-1}\left(\frac{R}{\sqrt{r^2 - R^2}}\right)\right]_{r=R}^\infty$$

$$= \left(1 - r_{\rm s}/R\right)^{-1/2} \pi \geqslant \pi \tag{4.21}$$

Therefore $|\Delta\phi| \geqslant \pi$ along a null geodesic in the $\theta = \frac{\pi}{2}$ plane. This was to be expected near a positive mass source (see figure 12).

Using this fact, we can connect any points on $\mathcal{I}^-$ and $\mathcal{I}^+$ which lie outside the intervals between the ends of this null geodesic and spatial infinity. The construction is illustrated in figure 13. In this figure we are interested in finding a timelike curve between the point $p \in \mathcal{I}^-$ (with $v = v_0$, $\phi = \phi_0$) and the point $q \in \mathcal{I}^+$ (with $u = u_1$, $\phi = \phi_1$). The details of this construction are as follows.

Suppose we are able to choose the impact parameter R such that the past and future endpoints of the null geodesic have $v = v_* > v_0$ and $u = u_* < u_1$ respectively. Along this geodesic, we
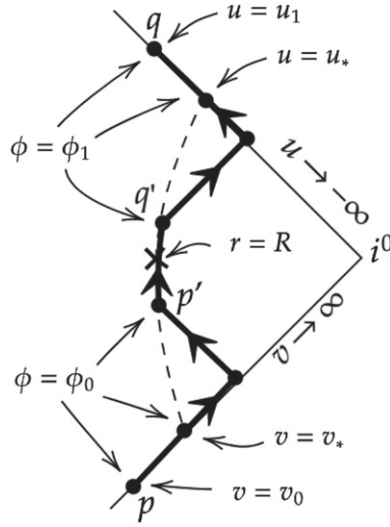
Figure 13. Construction used in [11] to timelike connect points $p \in \mathcal{I}^-$ and $q \in \mathcal{I}^+$. This is based on using a null geodesic (dashed line) to vary the angular co-ordinate ϕ (working in the hypersurface $\theta = \frac{\pi}{2}$). We then complete the path using null geodesics with zero angular momentum. The existence of a smooth timelike curve from p to q is then guaranteed by proposition 3.2.

choose points p' and q' (with $q' \in J^+(p')$) at which $\phi = \phi_0$ and $\phi = \phi_1$ respectively (we know this is possible because we have shown that $|\Delta\phi| \geqslant \pi$ along this null geodesic and we can choose the angular momentum h to be positive or negative). We use this segment of geodesic to vary the value of ϕ. We connect p to p' using two null geodesics with zero angular momentum (so ϕ is constant along them), one of which runs along null infinity towards i^0. We connect q to q' similarly. This gives us a piecewise null geodesic curve from p to q. We deduce, using repeated applications of Proposition 3.2, that a smooth timelike curve from p to q must exist. So we have reduced the Penrose property to a problem concerning the end points of null geodesics in Schwarzschild. The construction described above will allow us to connect all pairs of points on $\mathcal{I}^-$ and $\mathcal{I}^+$ if we are able to choose null geodesic endpoints arbitrarily close to spatial infinity. This also demonstrates how the Penrose property is a property of neighbourhoods of spatial infinity. If we know that we can connect all points arbitrarily close to i^0 then this construction shows that the Penrose property is satisfied.

Points on $\mathcal{I}^-$ and $\mathcal{I}^+$ can be labelled by co-ordinates (v, θ, ϕ) and (u, θ, ϕ) respectively. The u co-ordinate, u_*, at the future endpoint of a non-radial null geodesic, chosen to be symmetric about $t = 0$, is given in terms of the u co-ordinate at $r = R$, denoted u_R, by

$$u_* = u_R + \int_R^\infty \frac{du}{dr} \, dr = -R - r_s \, \log\left(\frac{R}{r_s} - 1\right) + \int_R^\infty \frac{du}{dr} \, dr \qquad (4.22)$$

In [11], Penrose shows that

$$\int_R^\infty \frac{du}{dr} \, dr = R + \mathcal{O}(1)$$

where $\mathcal{O}(1)$ denotes terms which are bounded as $R \to \infty$. This proves that $u_* \to -\infty$ as $R \to \infty$. A similar argument also shows that $v_* \to \infty$ as $R \to \infty$. This is exactly what we wanted. We can therefore use this construction to connect any pairs of points on $\mathcal{I}^-$ and $\mathcal{I}^+$ using a timelike curve. We simply need to use a geodesic with sufficiently large impact parameter. So we see that the Penrose property holds, and as a result the perturbative quantum gravity consistency condition $g < \eta$ described in section 2, is not satisfied. $\qquad\square$

4.2.2. Negative mass. We now consider the Schwarzschild metric in $3 + 1$ dimensions with negative mass. The compactified metric is now as in (4.19) except with

$$V(r) = 1 + \frac{r_s}{r}$$
$$\implies r_* = r - r_s \, \log(r/r_s + 1) \tag{4.23}$$

and the compactified manifold covers the whole spacetime $\{r > 0\}$ outside the naked singularity. This naked singularity will not be relevant for our discussion because, as explained in the positive mass case, we are concerned only with properties of the spacetime near i^0 (and hence at large values of r).

Proposition 4.3. *Negative mass Schwarzschild spacetime in $3 + 1$ dimensions does not satisfy the Penrose property.*

Proof. The compactified Schwarzschild line element (4.19) differs from the compactified Minkowski metric (3.7) only by the function $\frac{r^2}{V(r)r_*^2}$ which appears in front of the angular terms (this function is identically 1 for Minkowski). In fact, the similarity is deeper than this. Not only do the metrics look similar, we also note that the negative mass Schwarzschild spacetime is mapped to the same compactified manifold as the $r > 0$ portion of Minkowski. This means we have a bijection between curves in compactified Schwarzschild and curves in compactified Minkowski (with $r > 0$) which arises from identifying the co-ordinates (T, χ, θ, ϕ) defined on both these manifolds. It does not matter that these will correspond to different curves when we map back to the physically relevant uncompactified spacetimes.

To further this comparison, we need to understand the relation between the compactified metrics, which means looking more closely at the function $\frac{r^2}{V(r)r_*^2}$. We consider the large r asymptotics of this function. Expanding $V(r)$ and $r_*(r)$ to leading order, we have

$$\frac{r^2}{r_*^2 V(r)} = \left[1 + \frac{2r_s}{r} \, \log\left(r/r_s + 1\right) + O\left(\frac{r_s^2}{r^2}(\log\left(r/r_s + 1\right))^2\right) \right] \left[1 - \frac{r_s}{r} + O\left(\frac{r_s^2}{r^2}\right) \right]$$

$$= 1 + \left(2 \log\left(r/r_s + 1\right) - 1\right) \frac{r_s}{r} + O\left(\frac{r^2}{r_s^2}(\log\left(r/r_s + 1\right))^2\right)$$

$$> 1 \text{ for large } r \tag{4.24}$$

This means the conformally compactified negative mass Schwarzschild metric in $3 + 1$ dimensions can be bounded above by the conformally compactified Minkowski metric, i.e. $g < \eta$ (for large r).

We conclude from this that the Penrose property does not hold for negative mass in $3 + 1$ dimensional Schwarzschild spacetime. In particular, if we choose antipodal points on $\mathcal{I}^-$ and $\mathcal{I}^+$ sufficiently near i^0, then these cannot be timelike connected. To see this, suppose we

could connect such points using a Schwarzschild timelike curve. By choosing these points sufficiently close to i^0, we can ensure that such a curve is restricted to large enough r to ensure that inequality (4.24) holds. This is because curves near i^0 in compactified Schwarzschild are necessarily restricted to large values of r_* (since $(T, \chi) = (0, \pi)$ at i^0 with these co-ordinates defined as in equation (4.18)) and from equation (4.23) we see that $r \to \infty$ as $r_* \to \infty$. This curve would then also be timelike in compactified Minkowski spacetime and would connect the same points on $\mathcal{I}^\pm$. But we know (section 3) that antipodal points sufficiently near spatial infinity in Minkowski spacetime cannot be timelike connected, so we conclude (choosing our points even closer to i^0 if necessary) that the Penrose property does not hold for negative mass Schwarzschild in $3 + 1$ dimensions. □

This argument is similar to the argument used for negative mass in $2 + 1$ dimensions. The difference is that the constant $\alpha \in (0, 1)$ which created the deficit angle has now been replaced by the function $\frac{r^2}{V(r)r_*^2} \in (0, 1)$ which depends on r.

4.3. Schwarzschild spacetime in $d + 1$ dimensions with $d > 3$

4.3.1. Positive mass. It is tempting to try to use the same construction as [11] (and summarised in section 4.2.1) to show that the Penrose property is true for positive mass Schwarzschild in $3 + 1$ dimensions. In fact, this construction no longer works because the end-points of non-radial null geodesics do not approach i^0 as we let the impact parameter $R \to \infty$. In fact, the situation for positive mass Schwarzschild in higher dimensions is similar to the negative mass case in $3 + 1$ dimensions as we see below.

Proposition 4.4. *Positive mass Schwarzschild spacetime in $d + 1$ dimensions $(d > 3)$ does not satisfy the Penrose property.*

Proof. The Schwarzschild metric in $d + 1$ dimensions $(d > 3)$ is

$$\mathrm{d}s^2 = V(r)\mathrm{d}t^2 - \frac{\mathrm{d}r^2}{V(r)} - r^2 \ \mathrm{d}\omega_{d-1}^2 \tag{4.25}$$

where $V(r) = 1 - \frac{r_s^n}{r^n}$, $n = d - 2$ and $\mathrm{d}\omega_{d-1}^2$ is the round metric on S^{d-1}.

We can conformally compactify the metric as in lower dimensions to obtain the same form as equation (4.19) (with $\mathrm{d}\omega_2^2$ replaced by $\mathrm{d}\omega_{d-1}^2$ throughout). Once again we consider the large r asymptotics of the metric. We have

$$\mathrm{d}r_* = \frac{\mathrm{d}r}{V(r)} = \left(1 - \frac{r_s^n}{r^n}\right)^{-1}\mathrm{d}r = \left(1 + \frac{r_s^n}{r^n} + O\left(\frac{r_s^{2n}}{r^{2n}}\right)\right)\mathrm{d}r$$

$$\implies r_* = r\left(1 - \frac{r_s^n}{(n-1)r^n} + O\left(\frac{r_s^{2n}}{r^{2n}}\right)\right)$$

$$\implies \frac{r^2}{r_*^2 V(r)} = \left(1 + \frac{2r_s^n}{(n-1)r^n} + O\left(\frac{r_s^{2n}}{r^{2n}}\right)\right)\left(1 + \frac{r_s^n}{r^n} + O\left(\frac{r_s^{2n}}{r^{2n}}\right)\right) \tag{4.26}$$

$$= 1 + \frac{(n+1)r_s^n}{(n-1)r^n} + O\left(\frac{r_s^{2n}}{r^{2n}}\right)$$

$$> 1 \text{ for large } r.$$

So we obtain the same bounds as in section 4.2.2, and the same argument tells us that the Penrose property does not hold in this spacetime (again noting that curves near i^0 in Schwarzschild are restricted to large values of r_* and that $r \to \infty$ as $r_* \to \infty$). □

4.3.2. Negative mass. The negative mass Schwarzschild metric in spacetime of dimension $d + 1$ ($d > 3$) is as in equation (4.25) except with $V(r) = 1 + \frac{r_s^n}{r^n}$, ($n = d - 2$). The bound we obtained in (4.26) is now reversed, so we cannot use this method to conclude that the Penrose property is not satisfied. In fact, the lightcones of the naturally associated Minkowski metric are now contained inside the Schwarzschild lightcones (at least at large r). This suggests that the spacetime may have a chance of satisfying the Penrose property (we should at least be able to timelike connect all the same points as we could in Minkowski by using curves restricted to large r where this bound holds). However, as we let $r \to \infty$, the Schwarzschild and Minkowski lightcones converge towards each other. This means that we may be unable to connect antipodal points sufficiently near i^0, as is the case in Minkowski spacetime, since such curves would be restricted to increasingly large values of r. In fact, it turns out that the deflection of null geodesics away from the negative point mass (see figure 12) prevents us from timelike connecting certain pairs of antipodal points near i^0.

Theorem 4.5. *Negative mass Schwarzschild spacetime in $d + 1$ dimensions ($d > 3$) does not satisfy the Penrose property.*

To prove this theorem, we will require the following result:

Theorem 4.6 (Seifert [15], theorem 1). *Let $a, b \in M$ be such that the causal diamond $J^+(a) \cap J^-(b)$ is non-empty, compact and strongly causal. Then there exists a causal geodesic from a to b.*

Proof of theorem 4.5. Suppose we are aiming to timelike connect points $p \in \mathcal{I}^-$ and $q \in \mathcal{I}^+$. We choose co-ordinates $(t, r, \theta_1, \ldots, \theta_{d-2}, \phi)$ such that $\theta_1 = \ldots = \theta_{d-2} = \pi/2$ at p and q. Once again we can exploit the staticity and spherical symmetry of the metric and argue using lemma 3.1 that it suffices to consider curves restricted to the surface $\theta_1 = \ldots = \theta_{d-2} = \pi/2$. All curves and points referred to in this proof will be assumed to lie on this surface. A negative point mass will be repulsive and will deflect null geodesics away (figure 14). We therefore expect that, in contrast to positive mass, causal geodesics will undergo an angular change of strictly less than π. This is what we now prove. Solving the geodesic equations for the Schwarzschild metric, we find

$$r^2 \dot{\phi} = h = \text{constant}$$

$$\text{and } \dot{r}^2 = 1 - V(r) \left(k^2 + \frac{h^2}{r^2} \right) \tag{4.27}$$

where $k \geqslant 0$ is a constant such that $k = 0$ for a null geodesic and $k > 0$ for a timelike geodesic. We can determine h^2 in terms of R (the minimum value of r along the geodesic) by solving $\dot{r} = 0$ at $r = R$. We find that

$$h^2 = \frac{R^2}{V(R)} \left(1 - k^2 V(R) \right)$$

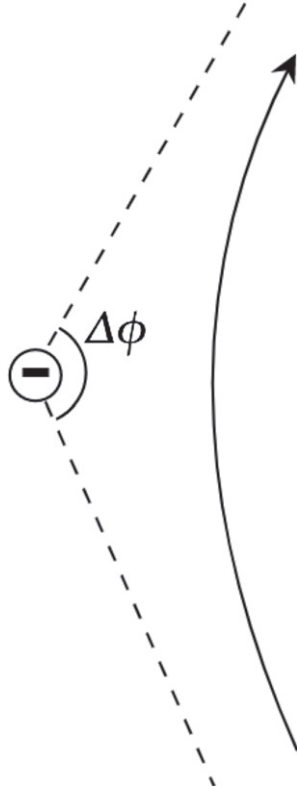$$\implies \frac{\dot{r}^2}{h^2} = \frac{V(R)}{R^2} \frac{1 - k^2 V(r)}{1 - k^2 V(R)} - \frac{V(r)}{r^2}$$

Figure 14. $|\Delta\phi| < \pi$ along null geodesics in negative mass Schwarzschild.

$$\geqslant \frac{V(R)}{R^2} - \frac{V(r)}{r^2}$$

$$\geqslant \frac{V(R)}{R^2}\left(1 - \frac{R^2}{r^2}\right) \tag{4.28}$$

where we have used the the fact that $V(r) \leqslant V(R)$ for $r \geqslant R$ and we also assume that $h \neq 0$ in order to make the inequalities easier to derive (the case $h = 0$ is trivial). This gives an upper bound for the magnitude of the change in ϕ along a causal geodesic, denoted $|\Delta\phi|$:

$$
\begin{aligned}
|\Delta\phi| &= 2|h|\int_R^\infty \frac{\mathrm{d}r}{\dot{r}r^2} \\
&\leqslant \frac{2R}{V(R)^{1/2}}\int_R^\infty \frac{\mathrm{d}r}{(1 - R^2/r^2)^{1/2}r^2} \\
&= \frac{\pi}{V(R)^{1/2}} \\
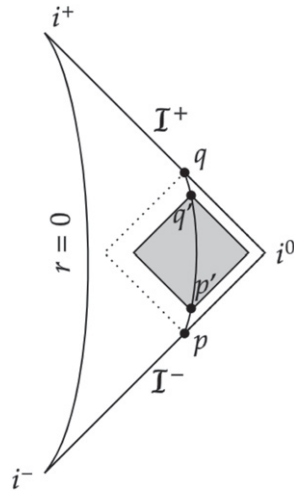&< \pi
\end{aligned}
\tag{4.29}
$$

Figure 15. The causal diamond (shaded) formed by the points p' and q' is non-empty, compact and strongly causal, so we can apply theorem 4.6.

where the final inequality is obtained from the fact that $V(R) = 1 + \frac{r_s^n}{R^n} > 1$. Note that the first inequality in (4.28) tells us that, as expected, among all causal geodesics with fixed h, $|\Delta\phi|$ is maximised by null geodesics (for which $k = 0$).

We aim to apply theorem 4.6 to the causal diamond formed by two points in the uncompactified spacetime (since inequality (4.29) refers to geodesics in this spacetime). Besides, since i^0 is singular, we could not apply the theorem to points on null infinity. In order to apply theorem 4.6, we need to make sure that the causal diamond does not intersect the naked singularity at $r = 0$.

Since the Penrose property relates to the compactified spacetime, we begin by choosing $p \in \mathcal{I}^-$ and $q \in \mathcal{I}^+$ sufficiently close to i^0 that $J^+(p) \cap J^-(q)$ does not intersect the $r = 0$ singularity. We can do this because the compactified spacetime can be decomposed as $\Sigma \times S^{d-1}$ as in lemma 3.1 . The causal diamond formed by any $p \in \mathcal{I}^-$ and $q \in \mathcal{I}^+$ then consists of a region on the $(d-1)$-sphere and a causal diamond on Σ (as shown by the dotted line in figure 15), which can certainly be chosen so as to avoid the $r = 0$ singularity.

Note in particular that the causal diamond formed by any two points in $J^+(p) \cap J^-(q)$ will also not intersect this singularity. Now if the Penrose property were satisfied in this spacetime then there would exist a timelike curve from p to q (recall that this curve would be timelike with respect to both the compactified and uncompactified metrics). Since the metric is invariant under $\phi \mapsto -\phi$, we can assume without loss of generality that ϕ is increasing along this curve. Now for any point $a \in \tilde{M}$, $I^\pm(a)$ are open sets ([12], proposition 2.8), so we must be able to modify this curve to obtain a new timelike curve, with the same u and v values at its endpoints, except with $\phi = -\epsilon$ on $\mathcal{I}^-$ and $\phi = \pi + \epsilon$ on $\mathcal{I}^+$ (for some $\epsilon > 0$) and such that ϕ is still increasing along this new curve. By restricting to a segment of this curve, we obtain a timelike curve between points p' and q' in the interior spacetime, where $\phi = 0$ at p' and $\phi = \pi$ at q'. The set $J^+(p') \cap J^-(q')$ (or more precisely, $\lambda^* \left(J^+(p') \cap J^-(q') \right)$, the pullback to the uncompactified manifold of this set by the embedding map defined in definition 2.2) then satisfies the conditions of theorem 4.6 (in particular it is free of any singular points). We therefore conclude that there must exist a causal geodesic (with respect to the uncompactified Schwarzschild metric) from $\lambda^*(p')$ to $\lambda^*(q')$. But we also know that such a geodesic cannot exist. In particular, it would necessarily have $\Delta\phi \geqslant \pi$ along

its full length which would contradict inequality (4.29). We therefore conclude that the Penrose property is not satisfied by negative mass Schwarzschild in higher dimensions.   □

Note that the spacetime is singular at i^0, so it was necessary to choose p' and q' to lie in the interior of the spacetime in order to ensure that $i^0 \notin J^+(p') \cap J^-(q)$ so we are able to apply theorem 4.6 to this region.

4.4. Summary of the Penrose property

Propositions 4.1, 4.3, 4.4 and theorems 4.2, 4.5 together establish theorem A from the Introduction which we now re-state.

Theorem A.   *The Penrose property is satisfied by Schwarzschild spacetime of mass m and varying spacetime dimension according to the following table*

Spacetime dimension	$m > 0$	$m \leqslant 0$
3	✓	✗
4	✓	✗
$\geqslant 5$	✗	✗

The Penrose property appears to be a property of positive mass in low dimensions. We have seen that adding positive mass in 3 or 4 dimensions gives us the inequality

$$\tilde{\mathrm{d}s}_S^2 \geqslant \tilde{\mathrm{d}s}_M^2 \tag{4.30}$$

at large r, where $\tilde{\mathrm{d}s}_M^2$ is the naturally associated compactified Minkowski metric. This tells us that the lightcone condition $g < \eta$ fails. This is also the case for negative mass in higher dimensions (although in this case, rather than diverging, the cones converge together at infinity and, as we have seen, the effect of the mass is not strong enough for the Penrose property to be satisfied). For negative mass in 3 and 4 dimensions as well as positive mass in higher dimensions, the inequality (4.30) is now reversed at large r and the Penrose property fails (by comparison with Minkowski space). It is curious that positive and negative mass seem to have the opposite effect on lightcones in higher dimensions compared with 3 and 4 dimensions.

Since the Penrose property is concerned only with the asymptotic behaviour of the metric near spatial infinity, these results can be immediately applied to a number of other spacetimes. For example the Reissner-Nordström line element in $d + 1$ dimensions ($d \geqslant 3$) is

$$\mathrm{d}s^2 = \left(1 - \frac{2m}{r^{d-2}} + \frac{Q^2}{r^{2d-4}}\right)\mathrm{d}t^2 - \left(1 - \frac{2m}{r^{d-2}} + \frac{Q^2}{r^{2d-4}}\right)^{-1}\mathrm{d}r^2 - r^2\,\mathrm{d}\omega_{d-1}^2$$

where Q is the charge of the black hole. Just like Schwarzschild, this spacetime satisfies the Penrose property only when $d = 3$ and $m > 0$. The case where $m = 0, Q \neq 0$ is notable because the large r asymptotics are now the same as those of negative mass Schwarzschild in $2d - 2$ spatial dimensions.

5. Kerr spacetime

We can also use this method of comparing metrics to gain insight into the Penrose property of a spacetime which is not spherically symmetric, namely the Kerr spacetime.

The Kerr metric in Boyer–Lindquist co-ordinates is

$$\mathrm{d}s_{K,m}^2 = \left(1 - \frac{2mr}{\Sigma}\right)\mathrm{d}t^2 - \frac{\Sigma}{\Delta_m}\mathrm{d}r^2 - \Sigma\mathrm{d}\theta^2 - \Lambda_m^2 \sin^2\theta \; \mathrm{d}\phi^2 + \frac{4mar \sin^2\theta}{\Sigma} \; \mathrm{d}\phi \; \mathrm{d}t$$

where we define

$$\Delta_m(r) = r^2 - 2mr + a^2$$
$$\Sigma(r,\theta) = r^2 + a^2 \cos^2\theta \tag{5.31}$$
$$\Lambda_m^2(r,\theta) = r^2 + a^2 + \frac{2ma^2 r \sin^2\theta}{\Sigma}$$

and $a = \frac{J}{m}$ is a constant rotation parameter written in terms of the angular momentum J and the mass $m > 0$ of the black hole. The subscripts m refer to the mass parameter and have been inserted for later convenience.

Theorem B. *The (positive mass) Kerr spacetime in $3 + 1$ dimensions satisfies the Penrose property.*

To prove this theorem, it will be useful to refer to the following metric which will be used to bound the Kerr line element.

Definition 5.1. For any $R > 2m > 0$, we define the *quasi-Schwarzschild* metric with parameters (m, a) by

$$\mathrm{d}s_{QS,m}^2 = \left(1 - \frac{2m}{r}\right)\mathrm{d}t^2 - \frac{\mathrm{d}r^2}{1 - \frac{2m}{r}} - r^2 \; \mathrm{d}\theta^2 - r^2 \sin^2\theta\bigl(\mathrm{d}\phi - \omega_{m,\alpha}(r,\theta)\;\mathrm{d}t\bigr)^2$$

where $\alpha^2 := 1 + \frac{a^2}{R^2}\left(1 + \frac{2m}{R}\right)$, $\omega_{m,\alpha}(r,\theta) := \frac{2mar\alpha^{-2}}{\Sigma(\frac{r}{\alpha},\theta)\Lambda_{\frac{m}{\alpha}}^2(\frac{r}{\alpha},\theta)}$. Again we have inserted subscripts for later convenience.

This metric is similar to the $3 + 1$ dimensional Schwarzschild metric of mass m, the only difference being the $\omega\mathrm{d}t$ term in the final bracket. Since we are interested in neighbourhoods of spatial infinity, it suffices to define the quasi-Schwarzschild metric on the manifold $M_R := \mathbb{R} \times \bigl(\mathbb{R}^3 \setminus B_R(0)\bigr)$, where $B_R(0)$ denotes the ball of radius R centred at the origin, for some $R > 2m$. For $r \geqslant R$, the components of the quasi-Schwarzschild metric are then smooth functions of the co-ordinates.

We can now prove the following lemma, which will be used in the proof of theorem B.

Lemma 5.2. *For $r \geqslant R$, the line element of the Kerr metric with mass m and rotation parameter a can be bounded from below by a quasi-Schwarzschild line element with parameters $(\tilde{m} = \alpha m, a)$:*

$$\mathrm{d}s_{K,m}^2 \geqslant \mathrm{d}s_{QS,\tilde{m}}^2$$

Proof.

$$
\begin{aligned}
\mathrm{d}s_{K,m}^2 &= \left(1 - \frac{2mr}{\Sigma} + \frac{4m^2a^2r^2\,\sin^2\theta}{\Sigma^2\Lambda_m^2}\right)\mathrm{d}t^2 - \frac{\Sigma}{\Delta_m}\,\mathrm{d}r^2 - \Sigma\mathrm{d}\theta^2 \\
&\quad - \Lambda_m^2\sin^2\theta\left(\mathrm{d}\phi - \frac{2mar}{\Sigma\Lambda_m^2}\mathrm{d}t\right)^2 \\
&\geqslant \left(1 - \frac{2m}{r}\right)\mathrm{d}t^2 - \frac{r^2+a^2}{r^2-2mr}\mathrm{d}r^2 - (r^2+a^2)\mathrm{d}\theta^2 \\
&\quad - \left(r^2 + a^2\left(1 + \frac{2m}{r}\right)\right)\sin^2\theta\big(\mathrm{d}\phi - \omega_{\alpha m,\alpha}(\alpha r,\theta)\mathrm{d}t\big)^2 \\
&\geqslant \left(1 - \frac{2\mathrm{m}}{r}\right)\mathrm{d}t^2 - \frac{\alpha^2\,\mathrm{d}r^2}{1 - \frac{2m}{r}} - \alpha^2 r^2\mathrm{d}\theta^2 \\
&\quad - \alpha^2 r^2\sin^2\theta\big(\mathrm{d}\phi - \omega_{\alpha m,\alpha}(\alpha r,\theta)\mathrm{d}t\big)^2 \\
&= \left(1 - \frac{2\tilde{m}}{\tilde{r}}\right)\mathrm{d}t^2 - \frac{\mathrm{d}\tilde{r}^2}{1 - \frac{2\tilde{m}}{\tilde{r}}} - \tilde{r}^2\,\mathrm{d}\theta^2 - \tilde{r}^2\,\sin^2\theta\big(\mathrm{d}\phi - \omega_{\tilde{m},\alpha}(\tilde{r},\theta)\mathrm{d}t\big)^2 \\
&= \mathrm{d}s_{QS,\tilde{m}}^2 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (5.32)
\end{aligned}
$$

where we have defined the new variable $\tilde{r} = \alpha r$. $\qquad\qquad\qquad\qquad\qquad\qquad\square$

From now on we will omit the subscripts α, m and $\tilde{m}$, with the understanding that for $r > R$, the Kerr spacetime of mass m should be compared to the quasi-Schwarzschild spacetime of mass $\tilde{m} = \alpha m$. The actual numerical value of $\tilde{m} > 0$ is unimportant. We simply require the existence of some quasi-Schwarzschild metric with positive mass with which to bound the Kerr metric.

Proof of Theorem B. We begin by compactifying the Kerr and quasi-Schwarzschild spacetimes. To compactify the quasi-Schwarzschild spacetime, we can define the same compactified co-ordinates and conformal factor as were used for Schwarzschild in section 4.2.1. This gives an embedding

$$\lambda_{QS} : (M_R, g_{QS}) \rightarrow (\tilde{M}_R, \tilde{g}_{QS})$$

from the quasi-Schwarzschild spacetime to its conformal compactification. The map λ_{QS} is invertible, so it is possible to express the $\omega_{m,\alpha}(r,\theta)\mathrm{d}t$ term of the quasi-Schwarzschild line element in terms of the compactified co-ordinates.

The compactification of the Kerr metric can be carried out by defining Kruskal co-ordinates as in section 8 of [14]. This gives an embedding

$$\lambda_K : (M_R, g_K) \rightarrow (\tilde{M}_R \tilde{g}_K)$$

We will use the fact that the compactifications of the Schwarzschild, quasi-Schwarzschild and Kerr metrics can all be defined on the same manifold $\tilde{M}_R$, so in particular we can identify points on $\mathcal{I}^\pm$ in these three spacetimes. This will be important in the argument below.

The compactified Kerr and compactified quasi-Schwarzschild line elements are conformal to the uncompactified versions. We have
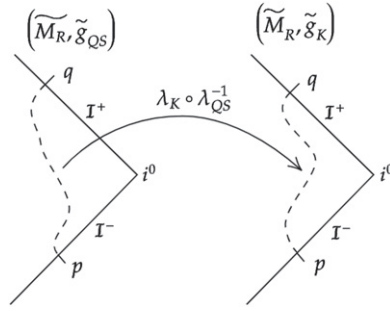
Figure 16. The function $\lambda_K \circ \lambda_{QS}^{-1}$ maps timelike curves in $(\tilde{M}_R, \tilde{g}_{QS})$ to timelike curves in $(\tilde{M}_R, \tilde{g}_K)$. This mapping also preserves the endpoints of these curves on $\mathcal{I}^{\pm}$.

$$\tilde{d}s_K^2 = \Omega_K^2 \ ds_K^2$$

$$\tilde{d}s_{QS}^2 = \Omega_{QS}^2 \ ds_{QS}^2$$

for some functions Ω_{QS} and Ω_K which in particular are strictly positive away from the boundary of the compactified spacetimes. From lemma 5.2 we have, for $r > R$:

$$ds_K^2 \geqslant ds_{QS}^2$$

$$\implies \Omega_K^{-2} \ \tilde{d}s_K^2 \geqslant \Omega_{QS}^{-2} \ \tilde{d}s_{QS}^2 \tag{5.33}$$

$$\implies \tilde{d}s_K^2 \geqslant \frac{\Omega_K^2}{\Omega_{QS}^2} \ \tilde{d}s_{QS}^2$$

So we have a lower bound for the compactified Kerr line element in terms of the compactified quasi-Schwarzschild line element. This bound tells us that the mapping $\lambda_K \circ \lambda_{QS}^{-1}$ maps timelike curves in compactified quasi-Schwarzschild spacetime to timelike curves in compactified Kerr spacetime.

Next we investigate how points on $\mathcal{I}^{\pm}$ behave under this mapping. It is possible that they relate to entirely different points on $\mathcal{I}^{\pm}$.[5] However, suppose we choose compactified co-ordinates defined by 'quasi-spherical light cones' as described in section 5 of [14]. This means compactifying the metric as in section 4.2.1, with r_* defined as in [14]. Then these compact-ified co-ordinates agree with the compactified Schwarzschild co-ordinates from section 4.2.1 as $r \to \infty$, so points on $\mathcal{I}^{\pm}$ are fixed under the mapping $\lambda_K \circ \lambda_{QS}^{-1}$. In particular, this means that if we have a timelike curve between two points on $\mathcal{I}^{\pm}$ in compactified quasi-Schwarzschild spacetime, then under the mapping $\lambda_K \circ \lambda_{QS}^{-1}$, this becomes a timelike curve between the same two points on $\mathcal{I}^{\pm}$ in the compactified Kerr spacetime figures.

Hence in order to show that the Penrose property holds in Kerr spacetime, it suffices to show that it holds in quasi-Schwarzschild spacetime. This is what we will now prove. For this argu-ment, we will refer to the Schwarzschild spacetime (M_R, g_S) in $3 + 1$ dimensions with mass

---

[5] This is precisely why this argument fails to show, for example, that positive mass Schwarzschild spacetime does not satisfy the Penrose property. In this situation we have $ds_S^2 \leqslant ds_M^2$, however when we compactify both metrics, the compactified co-ordinates diverge away from each other as we approach null infinity. This means that all timelike curves in compactified Schwarzschild correspond to curves between past and future timelike infinity in compactified Minkowski. As a result, we cannot use the failure of the Penrose property in Minkowski spacetime to deduce that it also fails in positive mass Schwarzschild.
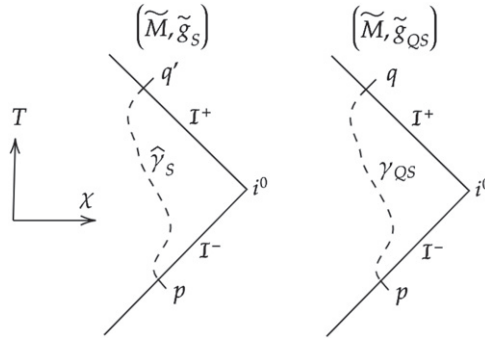
Figure 17. $\hat{\gamma}_S$ is a timelike curve in $(\tilde{M}_R, \tilde{g}_S)$ between p and q'. Using this curve, we define a curve γ_{QS} which is timelike in $(\tilde{M}_R, \tilde{g}_{QS})$ and which connects p and q. These curves follow the same path in (T, χ, θ) space.

$m > 0$. Suppose we wish to find a timelike curve in compactified quasi-Schwarzschild space-time connecting the point $p \in \mathcal{I}^-$ with co-ordinates $(T, \chi, \theta, \phi) = (T_0, \chi_0, \theta_0, 0)$ to the point $q \in \mathcal{I}^+$ with co-ordinates $(T_1, \chi_1, \theta_1, \phi_1)$ (we have exploited the axial symmetry of $(\tilde{M}_R, \tilde{g}_{QS})$ to set $\phi = 0$ at p). Since (M_R, g_S) satisfies the Penrose property, we can find a curve, γ_S, which is timelike with respect to g_S and which connects the point p to the point on $\mathcal{I}^+$ with co-ordinates $(T_1, \chi_1, \theta_1, \pi)$.

We can parametrize this curve by t and write the compactified co-ordinates along the curve as functions $T = T(t)$, $\chi = \chi(t)$, $\phi = \phi_S(t)$, $\theta = \theta(t)$. We have used a subscript S on the ϕ co-ordinate function to note that this refers to the Schwarzschild timelike curve γ_S. This is the function we will change when we construct a curve which is timelike with respect to the quasi-Schwarzschild metric.

We can also write $\omega = \omega(r(t), \theta(t))$ along this curve. This allows us to compute the following quantity:

$$A := \int_{t=-\infty}^{t=\infty} \omega(r(t), \theta(t)) \mathrm{d}t$$

Note that this integral converges since ω is finite at any r, θ (assuming $a \neq 0$) and for a path with endpoints on $\mathcal{I}^\pm$ we have

$$\omega \sim \frac{1}{r^3} \sim \frac{1}{r_*^3} \sim \frac{1}{|t|^3} \quad \text{as} \quad |t| \to \infty$$

where r_* is the co-ordinate defined in section 5 of [14]. The point about this definition is that $t \pm r_*$ tends to a constant near $\mathcal{I}^\pm$ and the leading order difference between r_* and r is a term which is logarithmic in r.

Observe that A is independent of the function $\phi_S(t)$. It depends only on the path in (T, χ, θ) space (since (t, r) are defined only in terms of (T, χ)).

Next note that if a curve is timelike in compactified Schwarzschild it will still be timelike if we reduce (and possibly reverse) the rate at which the ϕ co-ordinate varies, while keeping the same path in (T, χ, θ) space. By adding a suitable number of integer multiples of 2π to A if necessary, we may assume that $\phi_1 - A \in (-\pi, \pi]$. We then define a new curve

$$\hat{\gamma}_S(t) = \left(T(t), \chi(t), \frac{1}{\pi}(\phi_1 - A)\phi_S(t), \theta(t)\right)$$

This curve is the same as γ_S except with the ϕ co-ordinate function re-scaled by the constant $\frac{1}{\pi}(\phi_1 - A) \in (-1, 1]$, which in particular means that $\hat{\gamma}_S$ is timelike. It is important to note that A remains unchanged by this re-scaling. The reason for this re-scaling is that the endpoints of the curve $\hat{\gamma}_S$ satisfy

$$\hat{\phi}_S(t = -\infty) = 0$$
$$\hat{\phi}_S(t = \infty) = \phi_1 - A$$

(5.34)

We therefore have a curve, $\hat{\gamma}_S$, which is timelike with respect to g_S and which connects $p \in \mathcal{I}^-$ to the point $q' \in \mathcal{I}^+$ with co-ordinates $(T_1, \chi_1, \phi_1 - A, \theta_1)$.

Using this we can construct a curve, γ_{QS}, from p to q which is timelike with respect to g_{QS}. To do this, we define γ_{QS} to follow the same path in (T, χ, θ) space as the curve $\hat{\gamma}_S$ (figure 17). The ϕ co-ordinate along γ_{QS} will be given by a function, $\phi_{QS}(t)$, defined such that

$$\tilde{d}s_{QS}(t) \equiv \tilde{d}s_S(t) \geqslant 0$$

(5.35)

at every point on γ_{QS} and $\hat{\gamma}_S$. This ensures that γ_{QS} is timelike with respect to the quasi-Schwarzschild metric g_{QS}. Since $\hat{\gamma}_S$ and γ_{QS} follow the same path in (T, χ, θ) space (so the functions $r(t)$, $\theta(t)$ are the same along both curves), equation (5.35) becomes the requirement that

$$\dot{\phi}_{QS}(t) = \dot{\hat{\phi}}_S(t) + \omega(r(t), \theta(t))$$

To satisfy this, we define the function $\phi_{QS}(t)$ by

$$\phi_{QS}(t) = \hat{\phi}_S(t) + \int_{-\infty}^{t} \omega(r(t'), \theta(t')) dt'$$

This means that at the future endpoint of γ_{QS}, the ϕ co-ordinate is

$$\phi_{QS}(t = \infty) = \int_{-\infty}^{\infty} \dot{\phi}_{QS}(t) dt = \int_{-\infty}^{\infty} \left[\dot{\hat{\phi}}_S(t) + \omega(r(t), \theta(t)) \right] dt = \phi_0 - A + A = \phi_0$$

as required.　　　　　　　　　　　　　　　　　　　　　　　　　　　　□

6. Unique continuation of the linear wave equation

6.1. The null geodesic endpoint condition

The Penrose property, or more specifically the property of null geodesic endpoints proved by Penrose for positive mass Schwarzschild in $3 + 1$ dimensions [11], can help with understanding the problem of unique continuation from null infinity of the linear wave equation. In [1], the authors consider the linear wave equation

$$\Box_g \varphi + a^\mu \partial_\mu \varphi + V\varphi = 0$$

(6.36)

where $\Box_g$ is the Laplace–Beltrami operator for the metric g and the functions a^μ, V satisfy certain fall-off conditions (see [1]). They investigate conditions on the solution at null infinity which ensure that it can be uniquely continued into some region of the interior spacetime.

Two separate classes of spacetimes are considered. The first is Minkowski spacetime in $d + 1$ dimensions ($d \geqslant 2$) along with perturbations to the metric described in [1] (which are

such that the ADM mass [2] of the spacetime remains zero). The second class is Schwarzschild spacetime in $3 + 1$ dimensions along with metric perturbations [1]. Again these perturbations are sub-leading in the sense that they do not change the ADM mass of the spacetime. Note that this set-up is consistent with the proof in [11], summarised in section 4.2.1, that we cannot think of Schwarzschild spacetime in $3 + 1$ dimensions as arising from a perturbation of Minkowski due to differences in their asymptotic behaviours near i^0.

Following [1], we define the following subsets of null infinity

$$\mathcal{I}_{u_0}^+ = \{v = \infty, u \leqslant u_0\}, \quad \mathcal{I}_{v_0}^- = \{u = -\infty, v \geqslant -v_0\} \tag{6.37}$$

for any $u_0, v_0 \in \mathbb{R}$, where the retarded and advanced time co-ordinates u and v are given by equation (3.4) for perturbations of the Minkowski metric and equation (4.17) for perturbations of the Schwarzschild metric. For some $u_0, v_0 \in \mathbb{R}$, we will be interested in uniquely continuing the solution to equation (6.36) from $\mathcal{I}_{u_0}^+ \cup \mathcal{I}_{v_0}^-$ into a region

$$\mathcal{D}_\omega^{(u_0, v_0)} := \left\{ 0 < \frac{1}{(v + v_0)(u_0 - u)} < \omega \right\} \tag{6.38}$$

where $\omega > 0$.

It turns out that for perturbations of the Minkowski metric, if we impose conditions on the solution on subsets $\mathcal{I}_\epsilon^+ \subset \mathcal{I}^+$ and $\mathcal{I}_\epsilon^- \subset \mathcal{I}^-$ for any $\epsilon > 0$, then this solution can be uniquely continued into the interior spacetime . In particular, if we demand that φ and its first derivatives vanish to infinite order on $\mathcal{I}_\epsilon^+ \cup \mathcal{I}_\epsilon^-$ (for $\epsilon > 0$), then φ must vanish on some open domain in the interior spacetime which contains $\mathcal{I}_\epsilon^+ \cup \mathcal{I}_\epsilon^-$ on its boundary, i.e. $\varphi = 0$ on $\mathcal{D}_\omega^{(\epsilon, \epsilon)}$ for some $\omega > 0$

Now consider perturbations of the Schwarzschild metric in $3 + 1$ dimensions (which in particular include the Kerr spacetime). We define the sets $\mathcal{I}_{u_0}^+$, $\mathcal{I}_{v_0}^-$ and $\mathcal{D}_\omega^{(u_0, v_0)}$ as in (6.37) and (6.38), with the retarded and advanced time co-ordinates now defined by

$$u = t - r - r_{\mathrm{s}} \ \log(r/r_{\mathrm{s}} - 1)$$

$$v = t + r + r_{\mathrm{s}} \ \log(r/r_{\mathrm{s}} - 1)$$

Suppose we demand that the solution, φ, and its first derivatives vanish to infinite order on $\mathcal{I}_{u_0}^+ \cup \mathcal{I}_{v_0}^-$ for any $u_0, v_0 \in \mathbb{R}$. Then in [1] it is shown that φ must vanish in $\mathcal{D}_\omega^{(u_0, v_0)}$ for some $\omega > 0$.

Note that, in contrast to the case of perturbations of the Minkowski metric, we require conditions on φ and its first derivatives only on arbitrarily small sub-regions of $\mathcal{I}^+ \cup \mathcal{I}^-$ surrounding i^0 (since we are allowing $u_0, v_0 < 0$). This agrees with what we would have expected given the results obtained in previous sections. This is because for hyperbolic equations, initial data propagates along the characteristic curves. In this problem, the characteristic curves are the null curves of the metric g. Suppose that, having fixed φ and its first derivatives to vanish sufficiently quickly on some intervals $\mathcal{I}_{u_0}^+ \cup \mathcal{I}_{v_0}^-$, we are unable to extend the solution uniquely into a region $\mathcal{D}_\omega^{(u_0, v_0)}$ for any $\omega > 0$. Then for arbitrarily small $\omega > 0$, there must exist null curves which propagate from outside either of these intervals into $\mathcal{D}_\omega^{(u_0, v_0)}$. Equivalently, the following condition must fail to hold:

Null Geodesic Endpoint Condition: We say that a weakly asymptotically empty and simple spacetime, (M, g), satisfies the null geodesic endpoint condition if, given any $u_0, v_0 \in \mathbb{R}$, there exists $\omega > 0$ such that any inextendible null geodesic entering $\mathcal{D}_\omega^{(u_0, v_0)}$ must have at least one endpoint in $\mathcal{I}_{v_0}^-$ or $\mathcal{I}_{u_0}^+$.
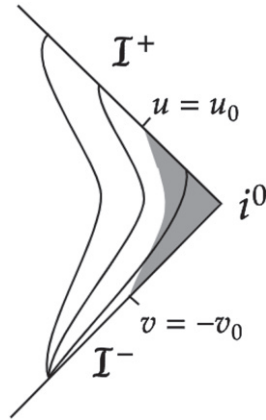
Figure 18. In Schwarzschild spacetime, requiring a null geodesic with past endpoint not in $\mathcal{I}_{v_0}^-$ to enter $\mathcal{D}_\omega^{(u_0,v_0)}$ (shaded region) for arbitrarily small $\omega > 0$ forces the future endpoint to lie arbitrarily close to i^0.

This is the property which Penrose shows [11] holds in positive mass Schwarzschild space-time in $3 + 1$ dimensions and which we have reviewed in section 4.2.1. Indeed for a null geodesic in this spacetime with fixed past endpoint not contained in $\mathcal{I}_{v_0}^-$, the requirement of entering $\mathcal{D}_\omega^{(u_0,v_0)}$ for arbitrarily small ω is equivalent to letting $R \to \infty$. This results in the future endpoint sliding along $\mathcal{I}^+$ towards i^0 and eventually entering $\mathcal{I}_{u_0}^+$ (see figure 18). This is a consequence of the stationarity of the spacetime, which allows us to extend the endpoint results of [11] and section 4.2.1. The null geodesics considered there were assumed to satisfy $r = R$ at $t = 0$ and hence were symmetric about $t \mapsto -t$. Translations in t will have the effect of moving the endpoints of these null geodesics along null infinity, so we can choose to keep the past endpoint fixed as we let $R \to \infty$, which we then see results in the future endpoint approaching i^0.

For Minkowski spacetime, this endpoint condition does not hold. We can, for example, fix the endpoints to lie at $u = 0$ and $v = 0$ while choosing the impact parameter to be arbitrarily large, so for any $u_0, v_0 < 0$, the null geodesic enters $\mathcal{D}_\omega^{(u_0,v_0)}$ for ω arbitrarily small (see figure 19). This is exactly why it is necessary to impose conditions on the solution on $\mathcal{I}_\epsilon^+ \cup \mathcal{I}_\epsilon^-$ for some $\epsilon > 0$ in order to achieve the unique continuation result. The solution will vanish in regions $\mathcal{D}_\omega^{(\epsilon,\epsilon)}$ which cannot be entered by any inextendible null geodesic with both endpoints outside $\mathcal{I}_\epsilon^+ \cup \mathcal{I}_\epsilon^-$.

6.2. General static, spherically symmetric spacetimes

We saw in section 4.3.1 that Penrose's method of proof does not extend to positive mass Schwarzschild in higher dimensions. This is because the null geodesic endpoint condition is no longer satisfied. In particular, if we impose $r = R$ at $t = 0$, then at the past and future endpoints we have $v \to 0^-$ and $u \to 0^+$ respectively as we let $R \to 0$. As noted in [1], this is why it is necessary to impose conditions on φ on $\mathcal{I}_\epsilon^+ \cup \mathcal{I}_\epsilon^-$ for some $\epsilon > 0$ (as in Minkowski) in order to obtain a unique continuation result.

As we have seen, to prove that the Penrose property holds for some spacetime we could try to use the same construction as Penrose in [11]. If the metric is stationary then, as discussed in the previous section, showing that non-radial null geodesic endpoints approach i^0
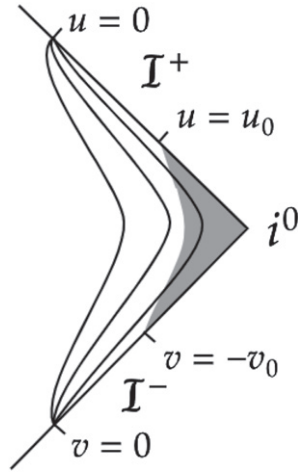
Figure 19. For any $u_0, v_0 < 0$, null geodesics in Minkowski spacetime with past and future endpoints lying outside $\mathcal{I}_{v_0}^-$ and $\mathcal{I}_{u_0}^+$ (e.g. at $v = 0$ and $u = 0$) can enter $\mathcal{D}_\omega^{(u_0,v_0)}$ (shaded region) for any $\omega > 0$.

as we let $R \to \infty$ is equivalent to proving that the null geodesic endpoint condition holds. If we now restrict attention further and consider only static, spherically symmetric spacetimes in $d + 1$ dimensions ($d \geqslant 3$) then the other result we require for the Penrose construction is that $|\Delta\phi| \geqslant \pi$ along all non-radial null geodesics (restricted to large r if necessary). A general static, spherically symmetric metric is characterised by two functions, $A(r)$ and $B(r)$, and in $d + 1$ dimensions can be written as

$$\mathrm{d}s^2 = A(r)^2\mathrm{d}t^2 - B(r)^2\mathrm{d}r^2 - r^2 \ \mathrm{d}\omega_{d-1}^2 \tag{6.39}$$

Since both the Penrose property and the null geodesic endpoint condition are conformally invariant, we are only interested in the conformal class of this metric. This allows us to remove one of these functional degrees of freedom. If we conformally re-scale the metric by a factor $A(r)^{-2}$, define a new radial co-ordinate $\tilde{r} = \frac{r}{A(r)}$ and then define the function[6] $C(\tilde{r}) = \frac{B(r(\tilde{r}))}{A(r(\tilde{r}))}\frac{\mathrm{d}r}{\mathrm{d}\tilde{r}}$, it suffices to consider a metric of the form (dropping tildes):

$$\mathrm{d}s^2 = \mathrm{d}t^2 - C(r)^2\mathrm{d}r^2 - r^2 \ \mathrm{d}\omega_{d-1}^2 \tag{6.40}$$

Using lemma 3.1, we can restrict attention to the $\Sigma \times S^1$ submanifold with co-ordinates (t, r, ϕ), on which the induced metric is

$$\mathrm{d}s^2 = \mathrm{d}t^2 - C(r)^2\mathrm{d}r^2 - r^2 \ \mathrm{d}\phi^2 \tag{6.41}$$

---

[6] We assume that we can invert the function $\tilde{r}(r)$ (restricted to large r if necessary). This is guaranteed to be the case if we assume, consistent with asymptotic flatness, that $A(r)$ can be expanded as a Taylor series in $1/r$ (as we assume for $C(r)$ in equation (6.42)). This ensures that $\tilde{r}(r)$ is a monotonic function (for sufficiently large r), and hence is invertible by the inverse function theorem. Recall that we can restrict attention to large r if necessary since the Penrose property is a property near i^0.

Consistent with asymptotic flatness, we assume we can expand:

$$C(r) = 1 + \frac{C_{d-2}}{r^{d-2}} + O\left(\frac{1}{r^{d-3}}\right) \tag{6.42}$$

A calculation similar to the one carried out in [11] shows that the null geodesic endpoint condition is satisfied if and only if $C_{d-2} > 0$ and $d = 3$. This is precisely the condition required to get the logarithmic divergence observed by Penrose. On the other hand, the angular condition ($|\Delta\phi| \geqslant \pi$ along non-radial null geodesics restricted to sufficiently large r) is satisfied if and only if $C(r) > 1$ for r sufficiently large. We therefore observe that this is a weaker condition than the null geodesic endpoint condition and hence that Penrose's construction works if and only if $C_{d-2} > 0$ and $d = 3$.

To determine if the Penrose property is satisfied for a metric of the form (6.40) we can use the same arguments as were used in previous sections for Schwarzschild spacetime. We have seen that if $d = 3$ and $C_1 > 0$ then the Penrose construction suffices to show that the Penrose property is satisfied. If $d = 3$ and $C_1 < 0$ or $d > 3$ and $C_{d-2} > 0$, then the comparison arguments outlined in sections 4.2.2 and 4.3.1 still hold and show that the Penrose property is not satisfied. If $C_{d-2} < 0$ and $d > 3$, we can use the same arguments as section 4.3.2 to again show that the Penrose property is not satisfied. The only remaining case is $C_{d-2} = 0$ ($d \geqslant 3$). This case is straightforward because either $C(r) \equiv 1$, in which case the metric is flat and we have seen that the Penrose property does not hold, or we have

$$C(r) = 1 + \frac{C_{d'-2}}{r^{d'-2}} + O\left(\frac{1}{r^{d'-3}}\right)$$

for some $d' > d \geqslant 3$, where $C_{d'-2} \neq 0$. In this case, the problem reduces to the question of whether the Penrose property is satisfied by some $d' + 1$ dimensional spacetime with metric of the form (6.40) [and inducing a metric of the form (6.41)], where $C_{d'-1} \neq 0$. But we have seen that the Penrose property is never satisfied for such higher dimensional spacetimes, so we conclude that for static, spherically symmetric spacetimes in $d + 1$ dimensions ($d \geqslant 3$), we have

$$\text{Penrose property} \iff [d = 3 \text{ and } C_1 > 0] \iff \text{Null Geodesic Endpoint Condition}$$

where C_1 is the constant arising from considering the conformally related metric of the form (6.40) and expanding $C(r)$ as in equation (6.42).

The constant C_{d-2} is equal to the ADM mass of the spacetime. However it does not make sense to phrase the Penrose property in terms of the ADM mass. This is because, unlike the Penrose property, the ADM mass (and even its sign) is not a conformally invariant quantity. In particular, in $3 + 1$ dimensions it is possible to have two metrics of the form (6.39) with the same functions $B(r)$ (and hence the same ADM mass) where one satisfies the Penrose property but the other does not (we can choose the function $A(r)$ such that the constant C_{d-2} has either sign).

Acknowledgments

ORCID iDs

Peter Cameron ⓘ https://orcid.org/0000-0002-5379-0179
Maciej Dunajski ⓘ https://orcid.org/0000-0002-6477-8319

References

[1] Alexakis S, Schlue V and Shao A 2016 Unique continuation from infinity for linear waves *Adv. Math.* **286** 481–544
[2] Arnowitt R, Deser S and Misner C 1962 The dynamics of general relativity *Gravitation: An Introduction to Current Research* ed L Witten (New York: Wiley)
[3] Camanho X O, Edelstein J D, Maldacena J and Zhiboedov A 2016 Causality constraints on corrections to the graviton three-point coupling *J. High Energy Phys.* JHEP02(2016)020
[4] Chruściel P and Galloway G 2004 A poor man's positive energy theorem *Class. Quantum Grav.* **21** L59–63
[5] Deser S, Jackiw R and 't Hooft G 1984 Three-dimensional Einstein gravity: dynamics of flat space *Ann. Phys., NY* **152** 220
[6] Dunajski M and Tod K P 2019 Conformally isometric embeddings and Hawking temperature *Class. Quantum Grav.* **36** 125005
[7] Galloway G 2002 Some global results for asymptotically simple space-times *The Conformal Structure of Space-Time* (*Lecture Notes in Physics* vol 64) (Berlin: Springer)
[8] Gao S and Wald R M 2000 Theorems on gravitational time delay and related issues *Class. Quantum Grav.* **17** 4999–5008
[9] Hawking S and Ellis G 1973 *The Large Scale Structure of Space-Time* (Cambridge: Cambridge University Press)
[10] Holland J and Sparling G 2018 Complex null geodesics in the extended Schwarzschild universe *Gen. Relat. Gravit.* **50** 86
[11] Penrose R 1980 On Schwarzschild causality-a problem for 'Lorentz covariant' general relativity *Essays in General Relativity: A. Taub Festschrift* ed F J Tipler (New York: Academic) pp 1–12
[12] Penrose R 1972 Society for industrial and applied mathematics *Techniques of Differential Topology in General Relativity* (Philadelphia, PA: SIAM)
[13] Penrose R, Sorkin R and Woolgar E 1993 A positive mass theorem based on the focusing and retardation of null geodesics (arXiv:gr-qc/9301015v2)
[14] Pretorius F and Israel W 1998 Quasi-spherical light cones of the Kerr geometry *Class. Quantum Grav.* **15** 2289
[15] Seifert H-J 1967 Global connectivity by timelike geodesics *Z. Naturforsch.* **22** 1356–60
[16] Staruszkiewicz A 1963 Gravitation theory in three-dimensional space *Acta Phys. Pol.* **24** 734
[17] Tod K P 2001 General relativity *Surveys in Differential Geometry* (*Essays on Einstein Manifolds* vol 6) ed C LeBrun and M Wang (International Press of Boston) pp 329–64
[18] Wald R 1984 *General Relativity* (Chicago, IL: University of Chicago Press)
[19] Witten E 2019 Light rays, singularities, and all that (arXiv:1901.03928.pdf)